OFFICE OF NAVAL RESEARCH

Grant N00014-92-J1052

R&T Code 4135018

Technical Report #8

"NMR Structural Studies on a Nonnatural Deoxyribonucleoside Mediated Recognition of GC Base Pairs in Pyrimidine•Purine•Pyrimidine Triplexes"

by
I. Radhakrishnan, D. J. Patel, E. S. Priestley, H. Nash, and P. B. Dervan

California Institute of Technology
Division of Chemistry and Chemical Engineering
Pasadena, CA

June 1, 1993

NMR Structural Studies on a Nonnatural Deoxyribonucleoside which Mediates Recognition of GC Base Pairs in Pyrimidine·Purine·Pyrimidine Triplexes.

Ishwar Radhakrishnan[a] and Dinshaw J. Patel[a,b]

[a]*Department of Biochemistry and Molecular Biophysics*
College of Physicians and Surgeons
Columbia University
New York, NY 10032

[b]*Cellular Biochemistry and Biophysics Program*
Memorial Sloan-Kettering Cancer Center
1275 York Avenue
New York, NY 10021

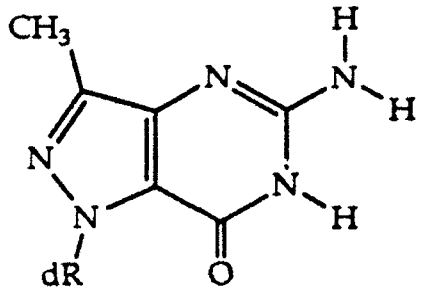
E. Scott Priestley,[c] Huw M. Nash,[c] and Peter B. Dervan[c]

[c]*Arnold and Mabel Beckman Laboratories of Chemical Synthesis*
California Institute of Technology
Pasadena, California 91125

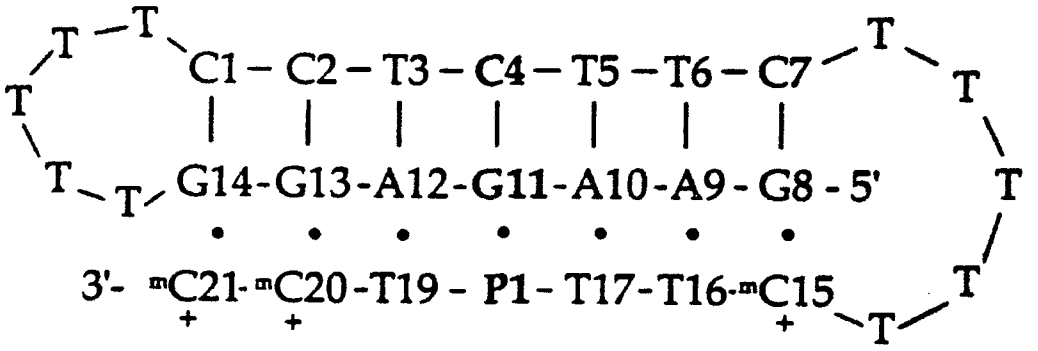
We are grateful to the Office of Naval Research for research support (PBD) and to the National Science Foundation for a predoctoral fellowship to ESP.

ABSTRACT

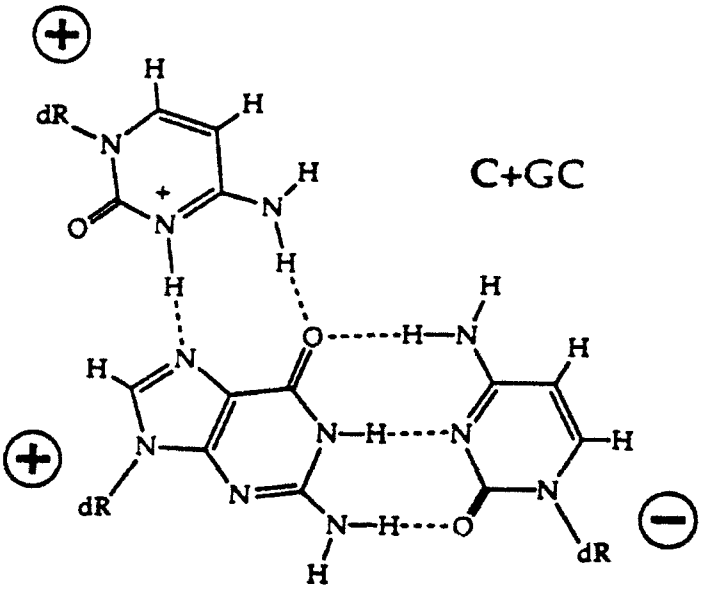
As a part of our ongoing efforts to define the structural aspects of unusual pairing alignments in DNA triplexes by Nuclear Magnetic Resonance spectroscopy, we have examined the structural role of a nonnatural deoxyribonucleoside, P1, that has been shown to mediate the recognition of GC base pairs in pyrimidine·purine·pyrimidine DNA triplexes (Koh, J. S. & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 1470). A qualitative interpretation of the NMR data indicates that this analog of protonated cytosines is readily accommodated in the third strand segment of an intramolecular triplex system. Furthermore, the observed NOE patterns position the imino and amino protons of P1 opposite the N^7 and O^6 atoms of guanine respectively, consistent with the previously proposed pairing scheme.



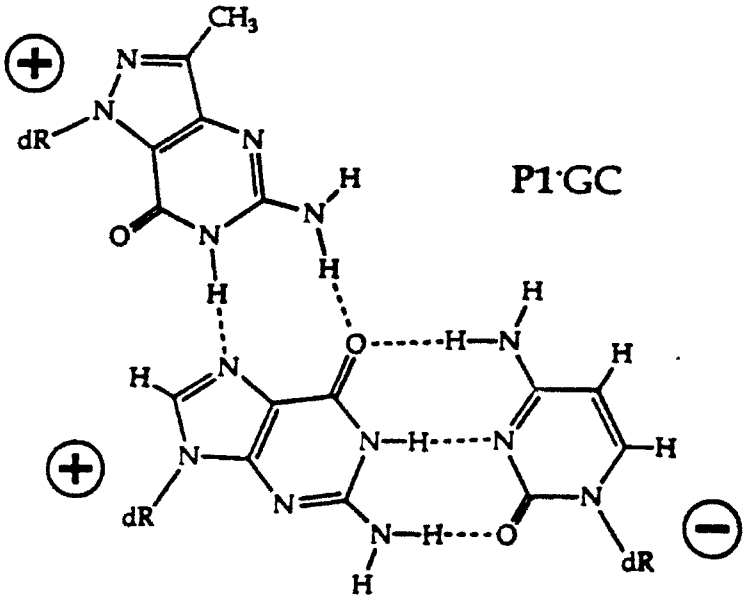
1



4



2



3

Table I

^{1}H chemical shifts for the exchangeable[a] and non-exchangeable[b] protons in the central tri-nucleotide segment in triplex 4.

Residue	H1/H3	H2[c]	H8/H6	NH_2[d]	H5/CH_3	H1'	H2', H2''	H3'
T3	14.39		7.63		1.78	6.04	*e* , 2.62	4.94
C4			7.69	8.75, 7.98	5.69	6.20	2.40, 2.76	4.80
T5	14.31		7.46		1.69	5.88	2.08, 2.60	4.93
A10		7.24	7.57	7.08, 7.13		5.99	2.32, 2.83	4.87
G11	12.62		6.81			5.85	2.32, 2.81	4.97
A12		7.45	7.25	7.70, 7.45		5.93	2.21, 2.80	4.84
T17	13.10		7.43		1.75	6.08	2.15, 2.84	4.97
P1	11.63[f]			7.07	2.22	6.45	2.63, 2.68	4.94
T18	12.88		7.40		0.93	6.06	2.32, 2.74	4.64

a: Chemical shifts for the exchangeable protons at 5 °C, pH 6.01 referenced to the residual solvent signal.

b: Chemical shifts for the non-exchangeable protons at 15 °C, pH 6.02 referenced to the residual HDO signal.

c: Chemical shifts for the H2 protons in H_2O at 5 °C, pH 6.01.

d: For cytidine and adenine amino protons, the chemical shift of the Watson-Crick hydrogen bonded proton is indicated first.

e: Could not be unambiguously assigned due to severe overlap.

f: Chemical shift is pH dependent.

FIGURE LEGENDS

Figure 1. The one-dimensional proton NMR spectrum of triplex **4** recorded in H_2O buffer at pH 6.01, 5 °C. (A) Only the region downfield to the H_2O resonance (6.5 to 15.5 ppm) is shown. The "jump-and-return" sequence (Plateau & Gueron, 1982) was used to suppress the H_2O signal with the delay between the two pulses set to 56 μs for maximum excitation of the hydrogen bonded imino proton resonances. (B) An expanded plot of the hydrogen bonded imino proton region (11.2 to 15.5 ppm). The assignments for the various imino proton resonances are labeled.

Figure 2. Duplicate, expanded contour plots of the symmetrical imino proton region (11.3 to 14.7 ppm) in the NOESY (mixing time 150 ms) spectrum of triplex **4** recorded in H_2O buffer, at pH 6.01, 5 °C. Hard 70° preparation and mixing pulses were used with the detection pulse replaced by a "jump-and-return" sequence (Plateau & Gueron, 1982). The carrier frequency was placed on the solvent signal and the spectral width was set to 11.1 kHz in both dimensions. The data were acquired with 1024 complex data points in t_2 and 400 complex data points in t_1 in the TPPI- States' mode (Marion, et al., 1989b). Prior to fourier transformation in t_2, the residual solvent signal in each increment was deconvoluted in the time domain (Marion et al., 1989a), apodized with a skewed 30° phase shifted sine bell window function and zero-filled to 2048 points. Prior to fourier transformation in t_1, baseline correction was applied in ω_2 and the interferograms apodized with a 90° phase shifted sine bell squared function and zero filled to yield a final processed matrix size of 2048 × 2048 in real points. (A) The lines trace the sequential connectivities between imino protons in adjacent base pairs for the duplex portion of triplex **4**. The tracing starts at G8 and proceeds via T6, T5, G11, T3, G13 all the way up to G14. (B) The tracing of sequential connectivities between adjacent bases in the third strand of triplex **4** starts at T16 and proceeds via T17 and P1 to T19. The N^3-protonated imino proton resonances of 5-methylcytosines in this strand exchange too rapidly with the solvent to show detectable NOEs.

Figure 3. Expanded contour plots of the same NOESY spectrum of triplex **4** shown in Figure 2. (A) NOE cross peaks correlating the imino protons (12.3 to 14.7 ppm) and the amino and base protons (6.4 to 9.2 ppm) in the Watson-Crick pairs in the triplex. The cross peak assignments for the protons in the central trinucleotide segment are as follows: *a*:T3(NH3)-A12(H2)/A12(NH_2-h); *b*:T3(NH3)-A12(NH_2-wc); *c*: T5(NH3)-A10(NH_2-wc)/A10(NH_2-h); *d*: T5(NH3)-A10(H2); *e*: G11(NH1)-C4(NH_2-e); *f*: G11(NH1)-A10(H2); *g*: G11(NH1)-A12(H2); *h*: G11(NH1)-C4(NH_2-b). (B) NOE cross peaks between the imino protons (11.3 to 14.7 ppm) and the amino and base protons (6.4 to 9.2 ppm) of the Hoogsteen pairs in the triples. The cross peak assignments are as follows: *a*: T16(NH3)-A9(NH_2-wc)/A9(NH_2-h); *b*: T16(NH3)-A9(H8); *c*: T16(NH3)-G8(H8); *d*: T17(NH3)-A10(NH_2-wc)/A10(NH_2-h); *e*: T17(NH3)-A10(H8); *f*: T17(NH3)-A9(H8); *g*: T19(NH3)-G11(H8); *h*: T19(NH3)-A12(H8); *i*: T19(NH3)-A12(NH_2-h); *j*: T19(NH3)-A12(NH_2-wc); k: P1(NH1)-G11(H8); l· P1(NH1)-(NH_2). In adenines, amino protons involved in Watson-Crick and Hoogsteen pairing are denoted by 'wc' and 'h' respectively while in cytosines, the hydrogen bonded and exposed amino protons are denoted by 'b' and 'e' respectively.

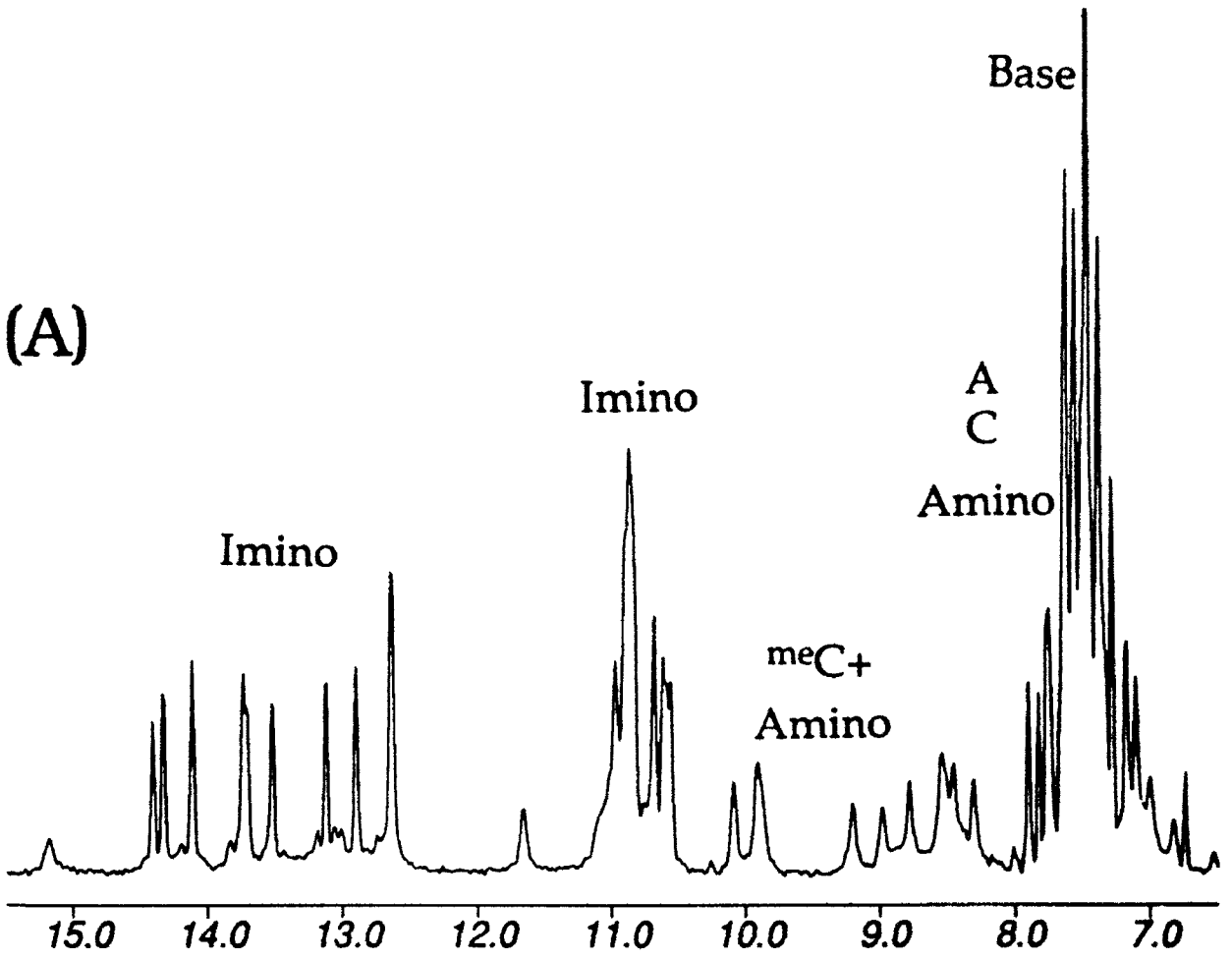
Figure 4. One-dimensional proton NMR spectra of a triplex **4** analog where cytosines have replaced 5-methylcytosines, recorded at 1° C in H_2O in the presence of 0.1 mM EDTA and 10 mM phosphate buffer at different pH values. Only the region downfield from the water resonance (6.5 to 16.0 ppm) have been plotted. The position of the P1 imino proton resonance in each spectrum is indicated by an asterisk.

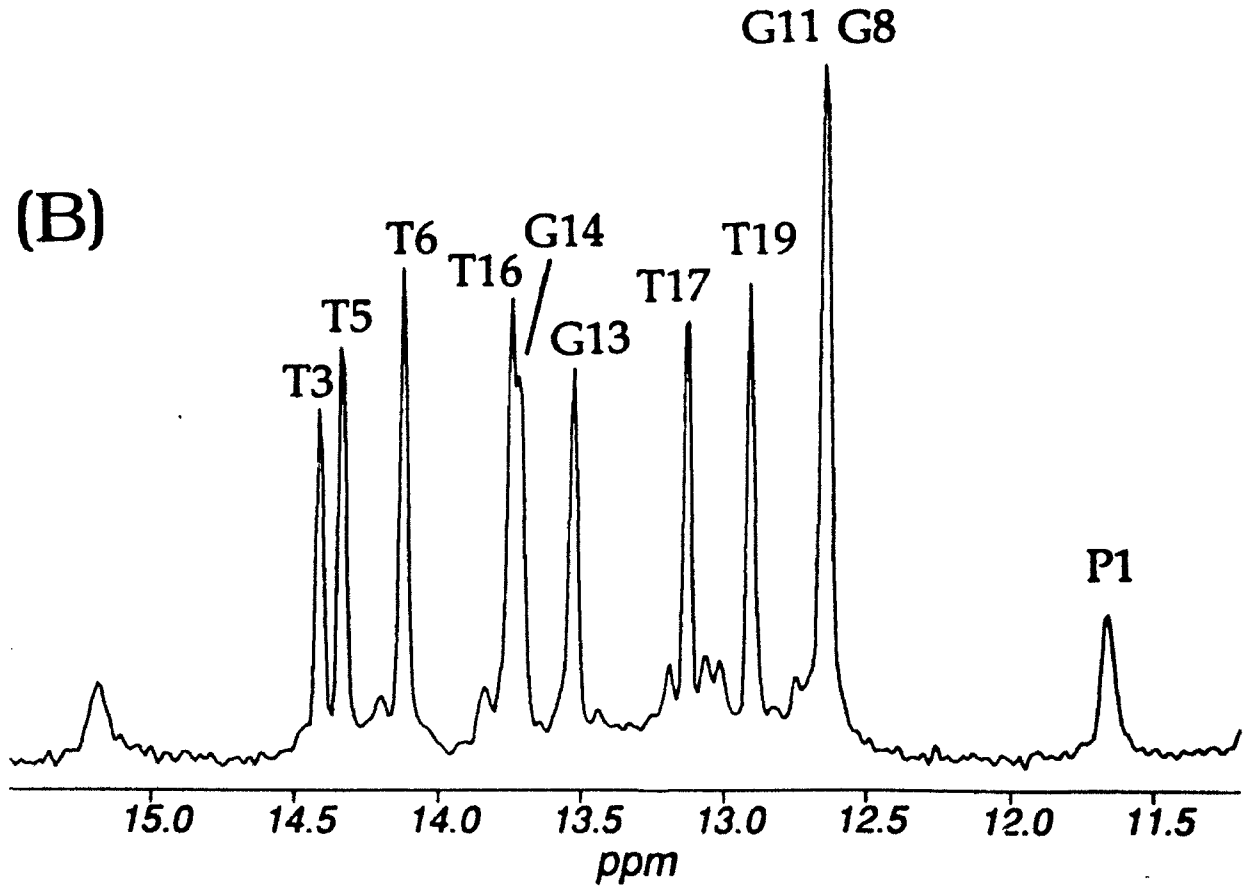
Figure 5. Plots of the chemical shifts of (A) the imino and (B) the methyl proton resonances of P1 as a function of pH in triplex **4** analog where cytosines have replaced 5-methylcytosines.

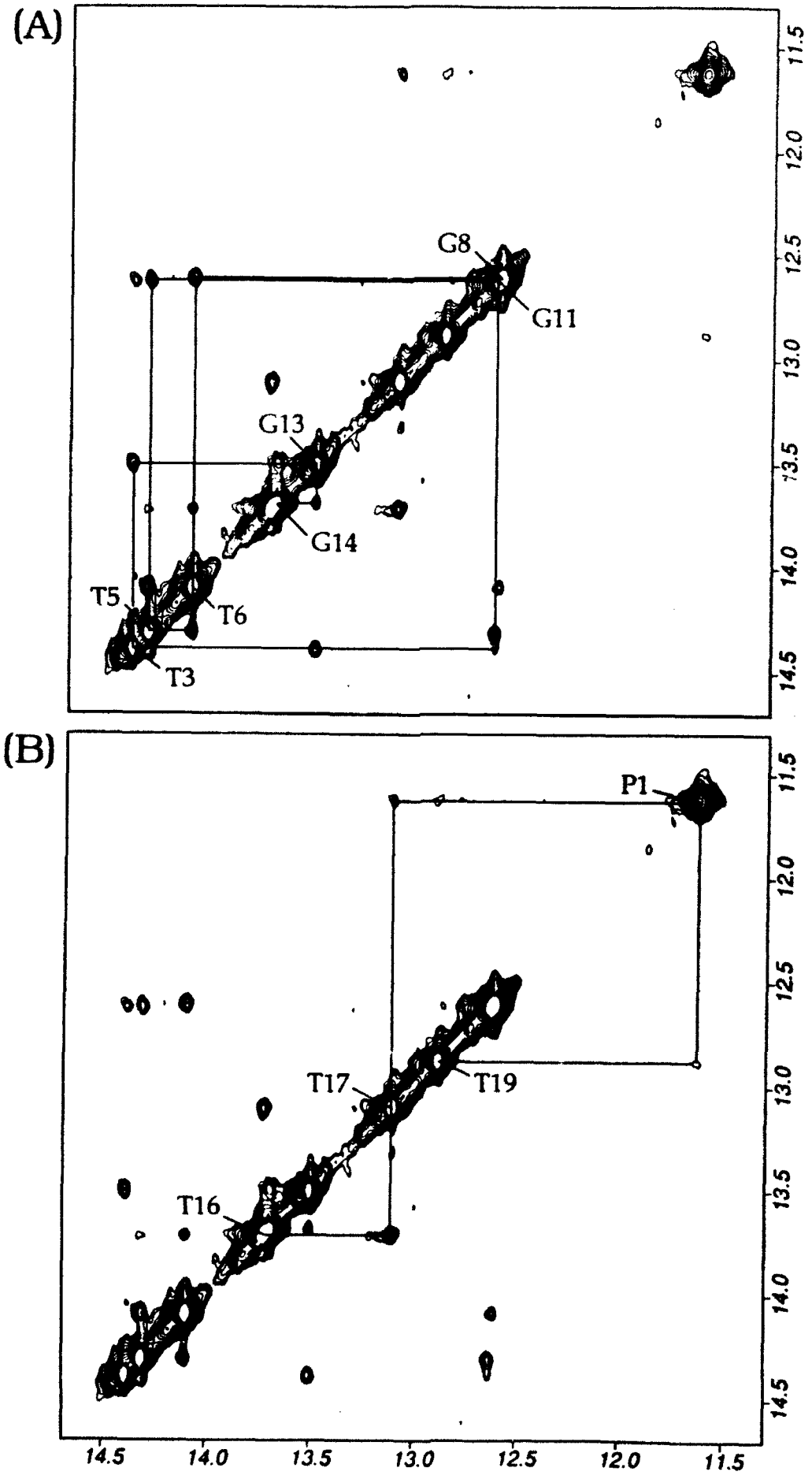
Figure 6. Duplicate, expanded contour plots correlating the base (6.7 to 8.1ppm) and sugar H1' protons (5.5 to 6.6 ppm) in the NOESY spectrum (mixing time 250 ms) of triplex **4** recorded in D_2O buffer, at pH 6.02, 15 °C. The data were acquired with 1024

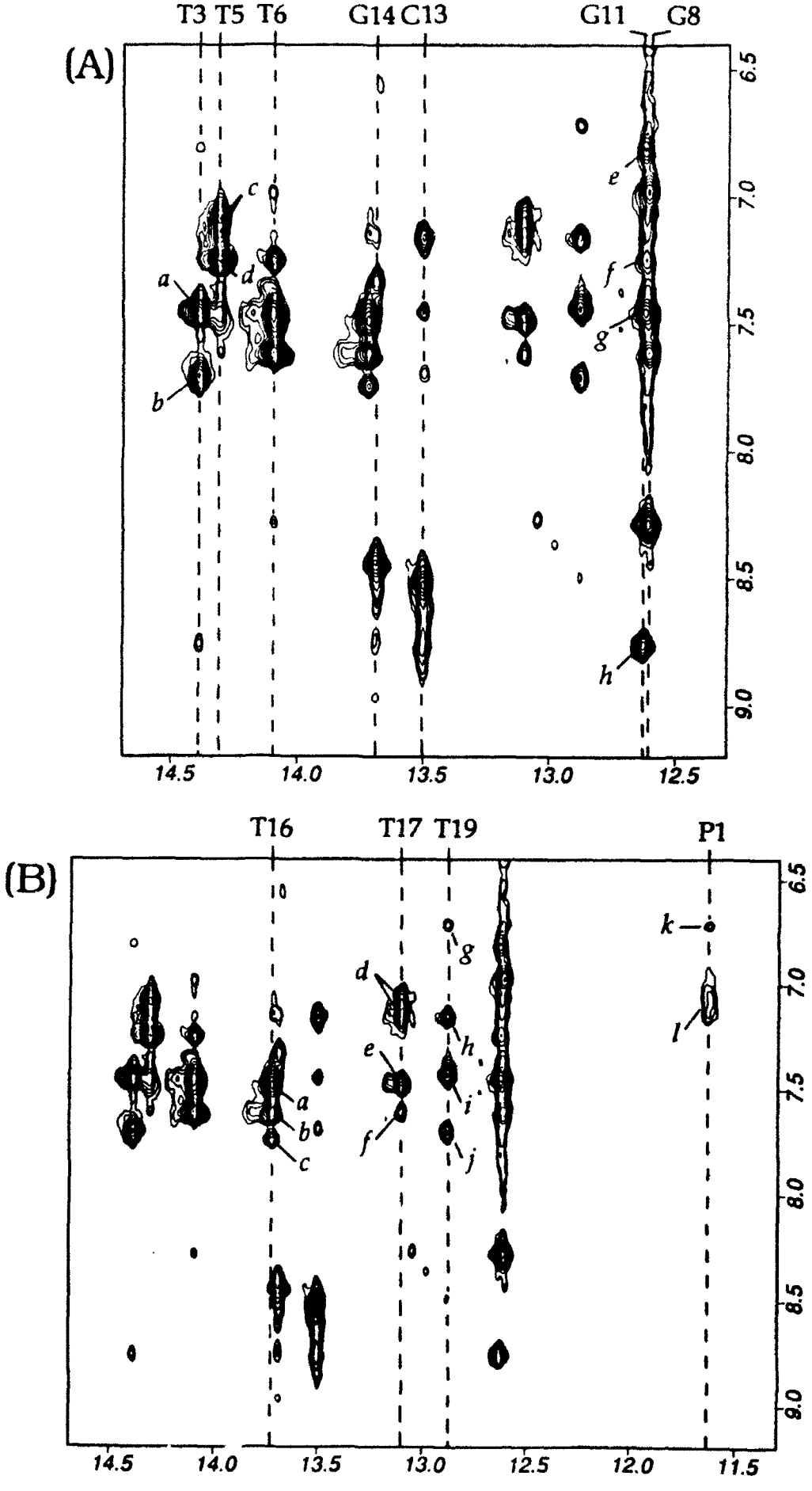
complex data points in t_2, 512 real data points in t_1 in the TPPI mode (Marion & Wüthrich, 1983) with the spectral width set to 5.05 kHz in both dimensions and a repetition delay of 1.6s. 90° phase shifted, sine beli squared and sine bell window functions were used for apodization in t_2 and t_1 dimensions, respectively. The data were zero-filled to 2048 points in both dimensions prior to fourier transformation. Baseline correction was applied to the frequency domain data in both dimensions. (A) The lines trace sequential NOE connectivities between the base and sugar H1' protons in the purine strand of triplex **4**. The tracing starts at G8 and traverses the strand up to G14. The intra-residue base to sugar H1' cross peaks in this strand are labeled by residue number. (B) The solid lines trace sequential NOE connectivities between the base and sugar H1' protons in the third strand of triplex **4** starting from ^{me}C15. The tracing continues up to the intra-residue base to sugar H1' cross peak of T17 where it is interrupted due to the absence of a base proton on P1 but resumes at the T19(H6)-P1(H1') cross peak and continues to ^{me}C21 at the end of the strand. The intra-residue cross peaks involving the base and sugar H1' protons in this strand are labeled by residue number. The dashed line corresponds to the position of the P1 H1' resonance. The cross-strand NOEs between the base protons in the purine strand and the H1' protons in the third strand are boxed and these are assigned as follows: *a*: G8(H8)-T16(H1'); *b*: A10(H8)-P1(H1'); *c*: G13(H8)-^{me}C21(H1'); *d*: A12(H8)-^{me}C20(H1'); *e*: G11(H8)-T19(H1').

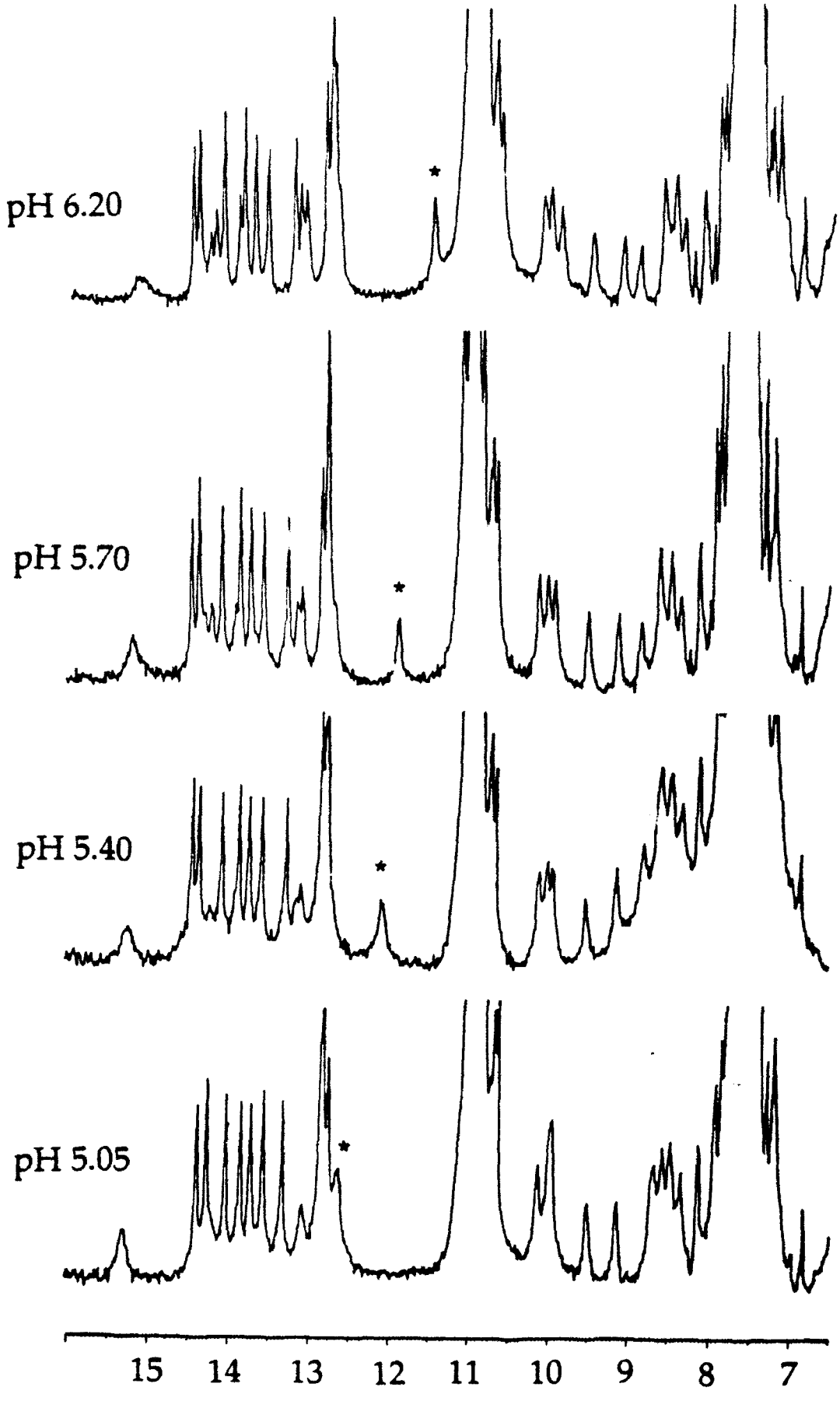
Figure 7. An expanded contour plot of the symmetrical methyl and H2', 2" proton region in the same NOESY spectrum of triplex **4** shown in Figure 6. The lines trace the sequential connectivities between methyl protons on adjacent bases in the third strand of the molecule starting from T17 and proceeding via P1, and T19 up to ^{me}C20. The boxed peak (indicated by an arrow) corresponds to the NOE observed between the sugar H2" proton of T17 and the methyl protons of P1.

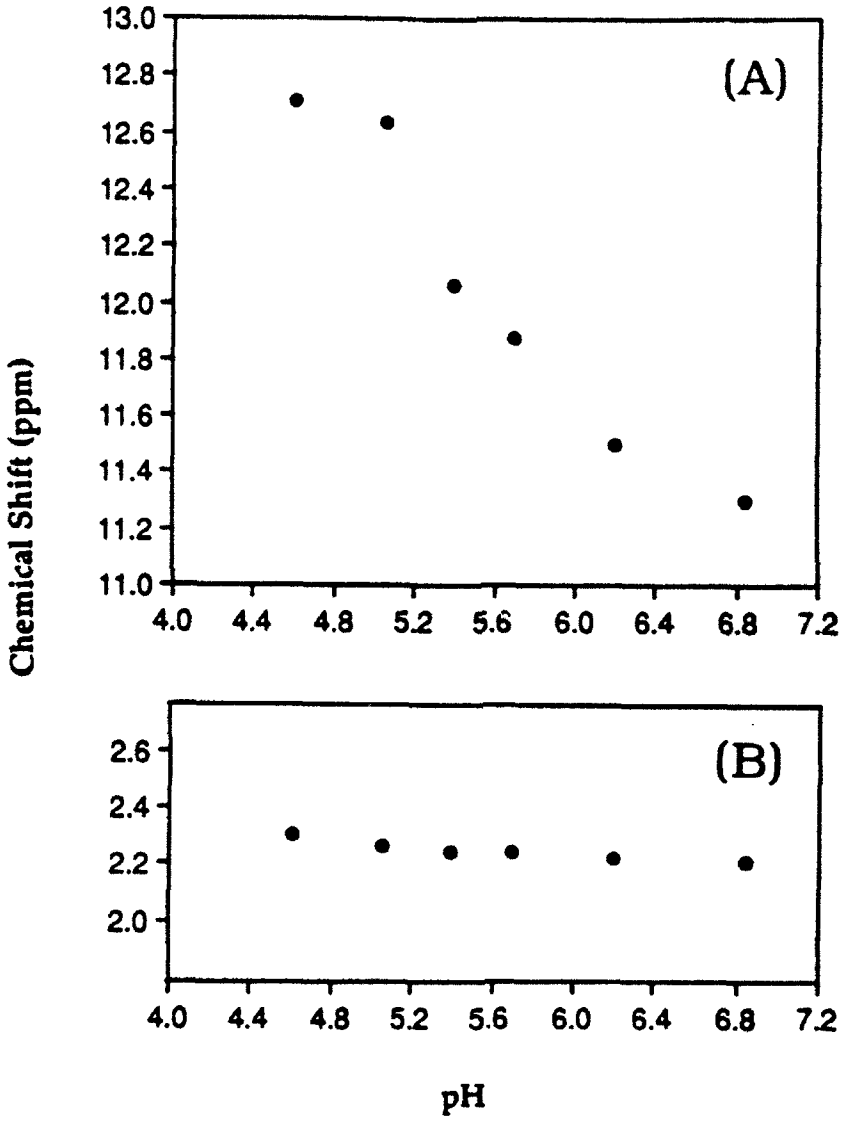

(A)
Base
Imino
A
C
Amino
Imino
meC+
Amino
15.0
14.0
13.0
12.0
11.0
10.0
9.0
8.0
7.0

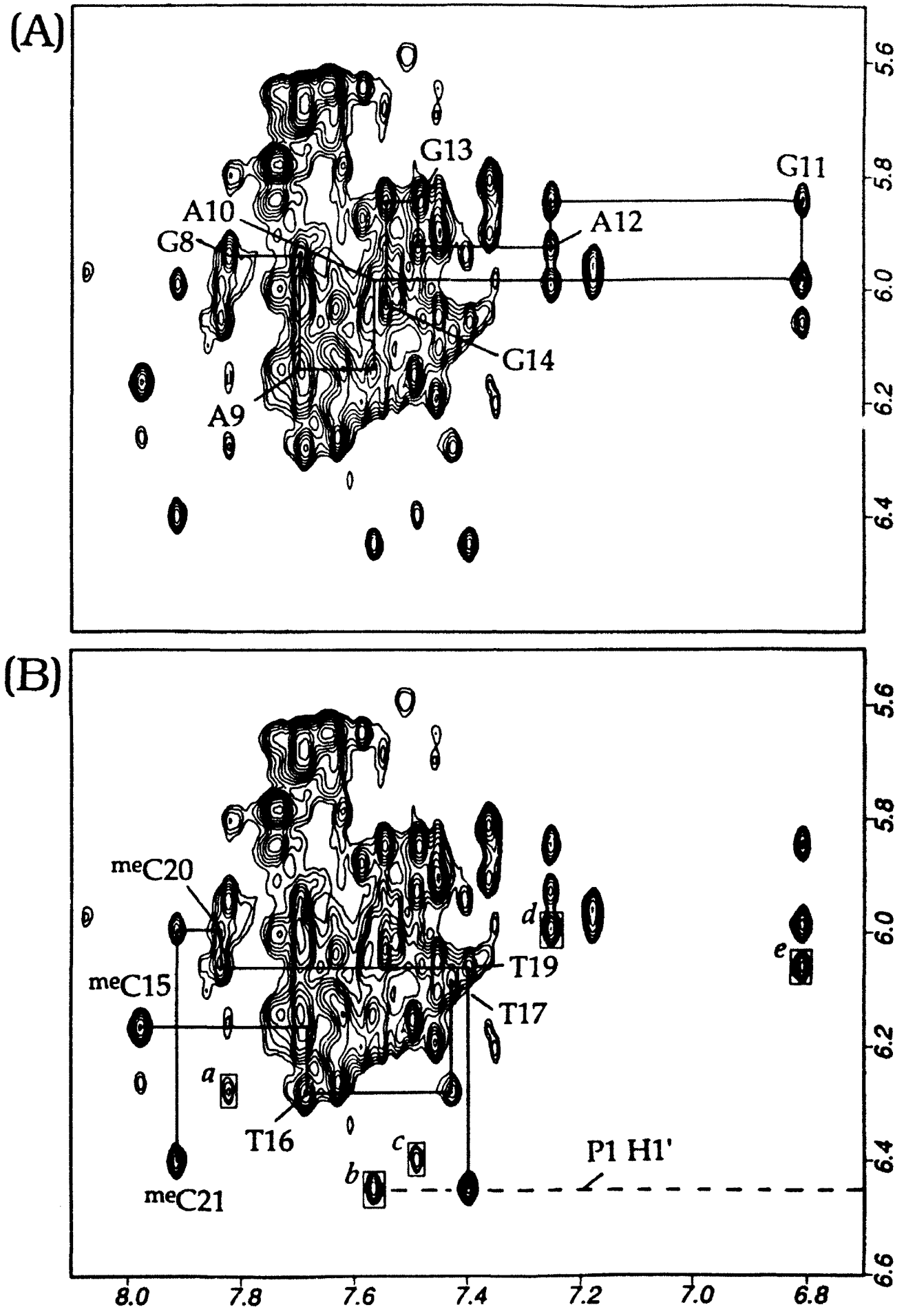

(B)
G11 G8
T6
G14
T16
T19
T5
T17
T3
G13
P1
15.0
14.5
14.0
13.5
13.0
12.5
12.0
11.5
ppm

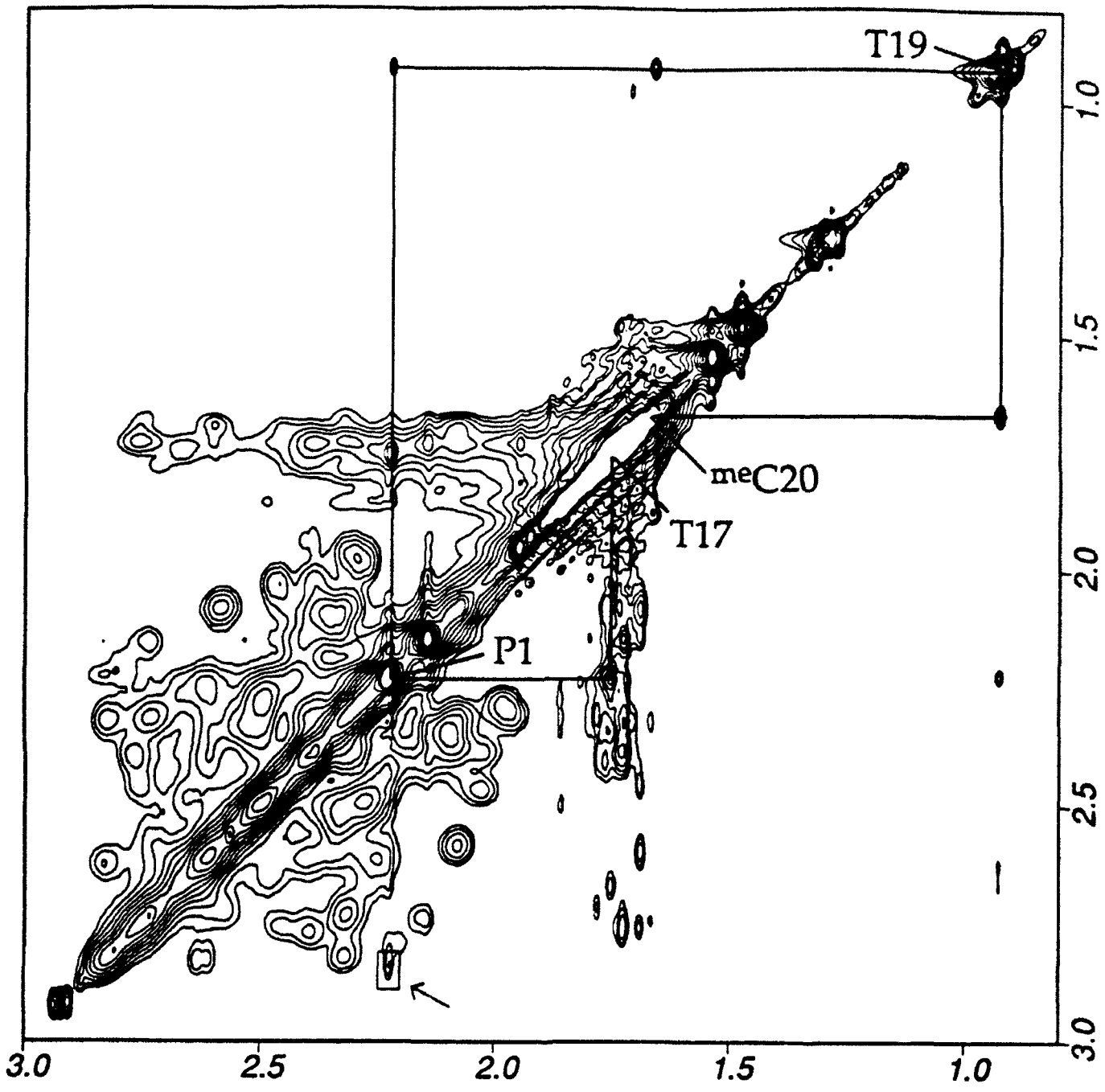

(A)
G8
G11
G13
G14
T5
T6
T3
11.5
12.0
12.5
13.0
13.5
14.0
14.5
(B)
P1
T17
T19
T16
14.5
14.0
13.5
13.0
12.5
12.0
11.5


(A)
T3 T5 T6
G14 C13
G11
G8
a
b
c
d
e
f
g
h
6.5
7.0
7.5
8.0
8.5
9.0
14.5
14.0
13.5
13.0
12.5
(B)
T16
T17 T19
P1
a
b
c
d
e
f
g
h
i
j
k
l
6.5
7.0
7.5
8.0
8.5
9.0
14.5
14.0
13.5
13.0
12.5
12.0
11.5


pH 6.20
pH 5.70
pH 5.40
pH 5.05
15
14
13
12
11
10
9
8
7


(A)
(B)
13.0
12.8
12.6
12.4
12.2
12.0
11.8
11.6
11.4
11.2
11.0
2.6
2.4
2.2
2.0
4.0
4.4
4.8
5.2
5.6
6.0
6.4
6.8
7.2
Chemical Shift (ppm)
pH


(A)
G13
G11
A10
A12
G8
G14
A9
5.6
5.8
6.0
6.2
6.4
(B)
meC20
d
e
T19
meC15
T17
a
T16
c
P1 H1'
b
meC21
5.6
5.8
6.0
6.2
6.4
6.6
8.0
7.8
7.6
7.4
7.2
7.0
6.8


T19
meC20
T17
P1
1.0
1.5
2.0
2.5
3.0
3.0
2.5
2.0
1.5
1.0

REFERENCES

Beal, P. A., & Dervan, P. B. (1991) *Science 251*, 1360-1363.

Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., & Hogan, M. E. (1988) *Science 241*, 456-459.

de los Santos, C., Rosen, M., & Patel, D. (1989) *Biochemistry 28*, 7282-7289.

Felsenfeld, G., Davies, D., & Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023-2024.

Gait, M. J. (1984) *Oligonucleotide synthesis: a practical approach*, IRL Press Limited, Oxford.

Koh, J. S., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 1470-1478.

Krawcyzk, S. H., Milligan, J. F., Wadwani, S., Moulds, C., Froehler, B. C., & Mateucci, M. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3761-3764.

LeDoan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J. L., Thoung, N. T., Lhomme, J., & Helene, C. (1987) *Nucleic Acids Res. 15*, 7749-7760.

Lin, C. H., & Patel, D. J. (1992) *J. Am. Chem. Soc. 114*, 10658-10660.

Maher, L. J., & Dervan, P. B. (1990) *Biochemistry 29*, 8820-8826.

Maher, L. J., Dervan, P. B., &Wold, B. (1992) *Biochemistry 31*, 70-81.

Marion, D., Ikura, M., & Bax, A. (1989a) *J. Magn. Reson. 84*, 425-430.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989b) *J. Magn. Reson. 85*, 393-399.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967-974.

Miller, P. S., Bahn, P., Cushman, C. D., & Trapane, T. L. (1992) *Biochemistry 31*, 6788-6793.

Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645-650.

Ono, A., Ts'o, P. O. P., & Kan, L.S. (1991) *J. Am. Chem. Soc. 113*, 4032-4033.

Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310-7311.

Plum, E. G., Park, Y.-W., Singleton, S. F., Dervan, P. B., & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9436-9440.

Povsic, T. J., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 3059-3061.

Radhakrishnan, I., Gao, X., de los Santos, C., Live, D., & Patel, D. J. (1991) *Biochemistry 30*, 9022-9030.

Radhakrishnan, I., Patel, D. J., & Gao, X. (1992) *Biochemistry 31*, 2514-2523.

Rajagopal, P., & Feigon, J. (1989) *Biochemistry 28*, 7859-7870.

Singleton, S. F., & Dervan, P. B. (1992) *Biochemistry 31*, 10995-11003.

Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Housman, D. E., & Dervan, P. B. (1991) *Science 254*, 1639-1642.

Watson, J. D., & Crick, F. H. C. (1953) *Nature 171*, 737-738.